融媒体背景下数据与业务结合的思考

摘　要：随着互联网的快速发展，大数据技术在各行各业中得到广泛应用。本文主要介绍大数据在公安媒体业务中的应用分析，以及大数据分析在公安媒体业务中的作用，让公安媒体业务充分发挥自有资源的价值优势，进一步提升公安新媒体运营和管理的水平，方便基层民警准确及时地处理事件，提升一线民警的战斗力和服务水平。

关键词：大数据；公安媒体；应用
中图分类号：G206.2　　文献标识码：A
文章编号：1671-0134（2019）08-035-03　　DOI：10.19483/j.cnki.11-4653/n.2019.08.009

文 / 赵振江

公安报业大数据，是把公安报历史资源、各业务系统生产资源、互联网数据、用户数据有机地整合、存储与管理在一起，为采编人员进行新闻素材获取、编辑、服务运营等业务工作提供共享资源池[1]。利用具有清晰分类导航的资源中心抓取文字稿库、图片稿库、新华社电稿、成品稿、微信、微博等各渠道资源，一键获取公安报内外部稿件资源，对各稿件的状态进行监控，提升公安报的数据分析与存储能力。智能分析的功能是以成熟的大数据分析为基础，利用人工智能技术进行深度学习，能够处理分析多种来源、多种形态数据，为策采编发、影响力评价、传媒及法律知识服务等业务开展及公安报新媒体运营和内容管理提供数据分析支撑，为公安机关提供互联网数据汇总分析、采编数据以及工具支持，方便基层民警使用，提高公安新闻传播力，提升一线民警的战斗力。[2]

公安报业大数据包含大数据采集、大数据管理、大数据分析、大数据应用四个层面。

1. 大数据采集

1.1 互联网数据采集

利用成熟的互联网信息采集工具以及数据中心实现实时监控和采集目标网站的内容，并把最新的网页及时采集到本地，进行内容分析和过滤等操作。[3] 采集工具的工作结果形成了新网页的全息信息集合，每个网页的详细信息被完整记录下来，包括网页名称、大小、日期、标题、文字内容、图片、表格等。

在实现全网数据自动化采集汇聚的同时，提供轻量型的指定内容一键获取工具。媒体人员在互联网上浏览到可能与公安报相关或写稿方向相关的内容时，通过一键获取工具，可以直接将内容推送至融媒体大数据云支撑系统，系统将自动完成页面内容的智能过滤、内容的结构化清洗等智能加工环节。

上述“人工＋智能化”的资源获取方式将全面而精确地锁定互联网数据源。

1.2 公安报内部数据资源整合

公安报内部资源整合主要包括从各个系统数据源中抽取数据、加工数据、存储数据。内部资源包括历史存量媒体资源、各渠道新生产资源以及第三方资源，主要包括文字、图片、图表、音频、视频、数字报刊、新媒体相关资源格式等，整合后的数据可以满足公安报社进一步数据挖掘需要。整合后的数据支持本地文件、共享文件、FTP 文件、HTTP 文件等常见数据源，也可根据具体要求扩展新数据源。[4]通过对内部数据源的整理和分析，便于进行内部数据资源的锁定与分析。

2. 公安报大数据的管理

大数据管理系统的存储与检索支撑融媒体内容生产、信息服务、运营服务平台业务的存储与检索服务。[5] 由于公安数据的特殊性，公安报大数据的管理，要具有高可靠的架构设计，需采用完全分布式的、多副本机制的、对等的、不共享的大数据管理系统，每新增加一个节点能同时增加系统的性能和存储容量，灵活实现对公安报增量数据的存储与管理。

3. 公安报大数据智能分析

公安报大数据智能分析是基于语义面向内容的智能化技术与大数据分析系统的融合，实现公安报数据服务的全方位智能支撑体系服务。

大数据智能分析系统通过充分利用大数据、云计算等前沿技术，结合自然语言处理技术与数据挖掘技术，融合中文信息处理、人工智能、信息检索的最新研究成果，以文本挖掘工具的自动分类、自动聚类等基本分析技术为基础，构建并融合管理大数据深度分析模型、警种专用数据分析模型，接收上层应用对工具、模型的调用，为各类应用提供技术支撑，大数据智能分析实现公

安报社资源、互联网资源价值的最大化。[6] 形成基于公安行业的大数据中心，搜集、汇聚、整理、分析和运用公安领域的内容大数据和用户大数据，提高新闻宣传实效性和满意度，促进媒体转型融合和增强舆论引导能力，为寻找新的服务模式奠定坚实的数据基础。

4. 大数据的应用

4.1 选题策划

选题策划需要抓住选题的时间深度、空间跨度、内涵深度、和传播广度。[7] 公安报基于大数据的云平台进行数据采集、挖掘和分析，对全网新闻信息进行智能化处理，构建新闻线索发现与分析、新闻热点发现与分析以及事件深度分析功能，为内部业务人员提供选题线索、选题方向指引。[8]

4.1.1 线索发现与分析

线索发现与分析利用线索探测聚类、相关性分析、重要性分析等算法实现全网实时线索聚焦、网民关注线索聚焦、微博爆料等功能，[9] 覆盖线索发现、线索聚合、线索监控、线索分析、线索日历等维度，帮助业务人员快速获取全网新闻线索，为报道选题策划提供支持。

4.1.2 素材发现与分析

素材发现与分析利用相关性分析等算法实现政策法规新闻、会议新闻、自然灾害新闻、安全事故新闻等信息的聚类、分类展示与分析，也可自主订阅素材。素材发现与分析利用相关信息发现关联，形成选题方向，同时为选题提供素材支持。

4.1.3 热点发现与分析

热点发现与分析利用热词挖掘、热点新闻评价、热点博客评价等技术，进行全网热点发现、地域热点发现、微博热点发现、头版头条聚类、排行榜单分析等，为选题策划提供数据支持。

4.1.4 热点事件深度分析

通过采用焦点还原、观点分析、趋势分析、情感分析等分析方法对热点事件进行多维度分析，使围绕该热点事件的深层次信息表现的更为简单直观。

4.2 公安新闻传播影响力

公安报利用融媒体大数据云支撑系统，获取原创稿件信息并进行全局统计分析。[10] 通过构建影响力分析模型和海量互联网指标数据，对全国媒体进行分级管理并获得传播力指数，形成覆盖“地域”“领域”“渠道”“作者”“部门”“栏目”“时间”等多维度的影响力分析，促进公安新闻整体传播力、影响力、公信力的提升。

4.2.1 稿件传播全局分析

以公安报全量稿件数据以及互联网传播数据为基础，统计分析相关稿件的发稿量、转载量和热议指数并做全局分析，对于热点稿件可以天为维度进行数量、变化趋势的可视化追踪。

4.2.2 稿件影响力

稿件影响力分析，具体包括影响力贡献分析、媒体采用影响力分析、稿件影响力渠道覆盖、稿件影响力地域分析、稿件影响力路径分析等。对公安报影响力稿件库内指定时间范围内的稿件可按照影响力指数进行排序。可根据稿件标题、作者、发稿日期、部门、分类、影响力指数、各渠道影响力指数等多维度进行解读。

4.3 公安传媒及法律知识

知识体系的建设和管理是公安传媒及法律知识服务的基础，利用知识深度加工的技术和方法，形成人民公安报特有的知识服务平台，包括法律知识库、典型案例剖析等，实现对政策法规、安全防范等知识资源的深加工，为公安媒体信息传播与应对提供重要的参考。

4.3.1 公安传媒分析

根据公安传媒的业务特点构建公安传媒资讯服务的多个维度的分类体系，以公安报融媒体的信息资源为基础，以公安传媒资讯的分类体系为依据，对相关事件进行分析，尤其是国家发布政法相关的最新政策，提供对最新政策的解读。

对于公安行业各子垂直领域的热点事件可进行专题分析。专题分析是对人民公安报社涉警舆情实时监测分析系统和公安警用大数据云支撑系统中大数据智能分析工具和分析模型的整合利用。可根据设定的条件自动汇聚与此专题相关的资源和信息，进行必要的统计分析和展现。例如，对“儿童拐卖”主题进行专题分析，可以形成如下形式的展现：汇集近年来儿童拐卖相关的新闻报道、破获的重大案件分析，以及不同渠道的舆论聚焦；儿童拐卖发生的地区形成可视化的地域分析；按时间维度对儿童拐卖案件数量形成趋势分析；将与儿童拐卖有关的法律法规及相关条款进行关联显示，如《未成年人保护法》《收养法》等。

4.3.2 法律法规知识

以中国特色社会主义法律体系和国家政策法规的基本分类，为法律法规知识库提供基础依据，并进行资源的梳理和展现，完成政策法规知识库的建设。法律法规知识库以完善、科学的知识分类为引导，为内外部用户提供法律法规条款条目的逐级浏览、查询服务。

知识库在对具体法律条款、政策法规进行展现的同时，也可与相关资源形成可视化的知识关联，包括典型案例剖析、政策解读、媒体报道等。

4.4 典型案例知识分析

建立典型案例库，从脉络分析、综合分析、舆论观点分析等角度完整刻画各个案例，围绕案例提供丰富的知识资源信息，实现对案例纵向和横向的全面了解。[11]

4.4.1 焦点脉络分析

以时间为线索，以案例分析时间范围和时间间隔为条件，同时，根据时间线上对案例发展具有重大影响的信息自动汇聚，结合一些重要时间节点及信息，形成针

对案例案件发生发展趋势的焦点脉络分析。

4.4.2 案例综合分析

可利用知识主题标引和关联分析等文本分析技术，结合知识服务平台中现有的知识资源，实现对案例的深度分析。具体可从媒体报道、涉案分析、专家解读和法律依据等方面进行分析。

4.4.3 舆论观点分析

围绕案例进行深度的舆论观点分析，对比分析官方媒体与民间媒体、社内媒体与社外媒体、境内媒体与境外媒体关于该案例产生的子话题，分析国内外重要人物、重要媒体、重要机构关于该事件所发表的观点评述。[12]

针对一些有争议的案例提供舆论观点的情感分析，对比分析各媒体及公众对案例中某些子话题的正面或负面观点。

4.4.4 同类案例推荐

利用关联分析、聚类分析等文本挖掘分析技术，在知识服务平台的案例库中查找与当前案例具有相似特点的案例进行关联推荐。相似性的分析可以从领域分类、领域主题、发生时期、发生地区、人员、法律法规等多个角度进行关联分析。

4.5 警种专用数据分析的模型

4.5.1 涉案人员分析模型

通过构建涉案人员分析模型，围绕案件嫌疑人及其相关人员，通过资源库导入人员档案信息，包括户籍、单位、住址等信息，也可以导入外部收集的数据，例如银行记录、话单等，在这些数据基础上，形成涉案人员在时间、空间等维度的关联关系分析，从而为案件的调查提供分析材料。

4.5.2 网上活动轨迹挖掘模型

构建重点人员网上活动轨迹挖掘指标体系与模型，为监测追踪重点人员的网上活动轨迹提供基础。以人物、事件为维度构建层级清晰的分析指标体系，形成单维度、组合维度的模型构建与可视化分析。

4.5.3 网络炒手分析模型

构建网络炒手分析指标体系与模型，为网络炒手的多维度分析提供基础。其指标体系由多层级指标构成，包括一级指标发布人基本信息、二级指标网络关系特征，可包括网民、所属媒体或组织、博客、微博、微博的好友、话题讨论关系。在词基础上，构建发布人特征分析、关联分析、影响力分析等模型基础，综合形成网络炒手分析模型。

4.5.4 情报预警模型

构建情报预警指标体系与模型，快速发现互联网情报信息。情报预警指标包括互联网情报点击量、回帖量的异常暴增、预设关键词、重点网民内容异常、重点网站内容异常等，通过赋予不同指标权重，通过预设阈值出发预警信号。

4.5.5 专项模型

为建立警务人员和各平台之间的联系，根据地方公安机关的具体需求，对警务人员属性信息、生活信息、工作信息的分析设立专项模型。

结语

人民公安报社充分运用大数据的分析方法进行数据挖掘、智能信息处理、知识发现与管理等先进技术，[13] 针对已发布的公安法规、政策、规划及落实情况、行业数据等信息，以及国家重大战略规划和产业调整，财税、环保、国土、能源等相关行业政策等信息进行深度挖掘和分析，构建服务于公安全链条的多维度的知识库，增强对政策评估和理论研究的能力，提升行业治理能力，善于在多元意见中占据主动性，争取引导舆论走向，进一步提升公安新媒体运营和管理的水平，方便基层民警准确及时地处理事件，提升一线民警的战斗力和服务水平。媒

参考文献

[1] 周君 . 警用信息系统的大数据技术应用 [J]. 产业与科技论坛，2015（4）：64–65.

[2] 李健财 . 大数据背景下智慧公安的发展策略 [J]. 电子技术与软件工程，2019（8）：157.

[3] 彭知辉 . 大数据：开启公安情报工作新时代 [J]. 公安研究，2014（1）：76–80.

[4] 张兆端 .“智慧警务”：大数据时代的警务模式 [J]. 公安研究，2014（6）：19–26.

[5] 彭知辉 . 论基于大数据的公安情报搜集 [J]. 图书馆学研究，2017（9）：33–37.

[6] 彭知辉 . 论大数据环境下公安情报流程的优化 [J]. 情报杂志，2016，35（4）：15–20.

[7] 吴峰平 . 重大主题报道的选题策划 [J]. 中国广播，2016(3)：70–71.

[8] 刘桂玲 . “互联网 +”视域下的公安政务双微研究 [J]. 山东警察学院学报，2016，28（4）：154–160.

[9] 王懂懂，魏永忠 . 大数据 3.0 时代：智慧警务建设的实践与思考 [J]. 广西警察学院学报，2018，31（1）：65–70.

[10] 刘旸 . 对公安机关提升新媒体沟通能力的思考 [J]. 现代交际，2014（3）：85–86.

[11] 房汉平 . 基于大数据的公安警务实训共享平台构建与运用 [J]. 实验技术与管理，2018，35（11）：171–175.

[12] 卿立新 . 突发公共事件网络舆论及其应对研究 [D]. 湖南师范大学，2013.

[13] 王崇 . 公安机关应对网络舆情危机的对策研究 [D]. 天津大学，2017.

（作者单位：人民公安报社）